METHODOLOGICAL AND LOGISTICAL ISSUES

IN

SOCIAL EXPERIMENTS:

THE CASE OF MINCOME

by

Greg Mason

Institute for Social and Economic Research

Faculty of Arts
The University of Manitoba
Winnipeg, Manitoba R3T 2N2

METHODOLOGICAL AND LOGISTICAL ISSUES

IN

SOCIAL EXPERIMENTS:

THE CASE OF MINCOME

by

Greg Mason

MINCOME MANITOBA

Working Paper No. 1

1.0 INTRODUCTION

The Manitoba Basic Annual Income Experiment (MINCOME) was a major social experiment patterned after income maintenance experiments in the United States. Jointly sponsored by the governments of Canada and Manitoba, it was designed to evaluate how family units reacted to a negative income tax, primarily with respect to labour supply, but also with a wide variety of other outcomes. Commencing in 1974, the experiment was terminated in 1979 prior to any substantive tests of the hypothesis having been done.

MINCOME is representative of many other large scale evaluative exercises undertaken in North America over the past decade. Thus far, these experiments have failed to substantially persuade policy makers, the concensus now appears to be that they failed in their mission. Does the fact that no more elaborate experiments are planned mean the end to the "experimenting societies" as envisioned by Campbell (1971), or have they suffered the same general fate of much social policy namely a fiscal crisis coupled with a more conservative approach to policy design?

This paper examines MINCOME as an example of social experimentation and contends that a number of methodological and logistical issues produced a degree of complication beyond the technological capability of the social science of everyday. These problems appear to have characterized to one degree or another all the social experiements attempted in the last while. No discussion is attempted on the bureaucratic and political complications which confronted MINCOME as these tend to be specific to institutions and personalities not likely duplicated in other situations. Despite the manifold problems faced by MINCOME and the delay in producing an accessible secondary data base, the information is extremely robust and valuable. Whatever the methodological and logistical problems may have been, they seem to not have compromised the information and its value as a secondary data base.

2.0 AN OVERVIEW OF MINCOME

It is important to overview negative income tax experiments to properly understand MINCOME. Unavoidably, there is a certain amount of jargon specific to income maintenance experiments and to MINCOME which must be mastered before a clear idea of the essential methodological and logistical problems is formed.
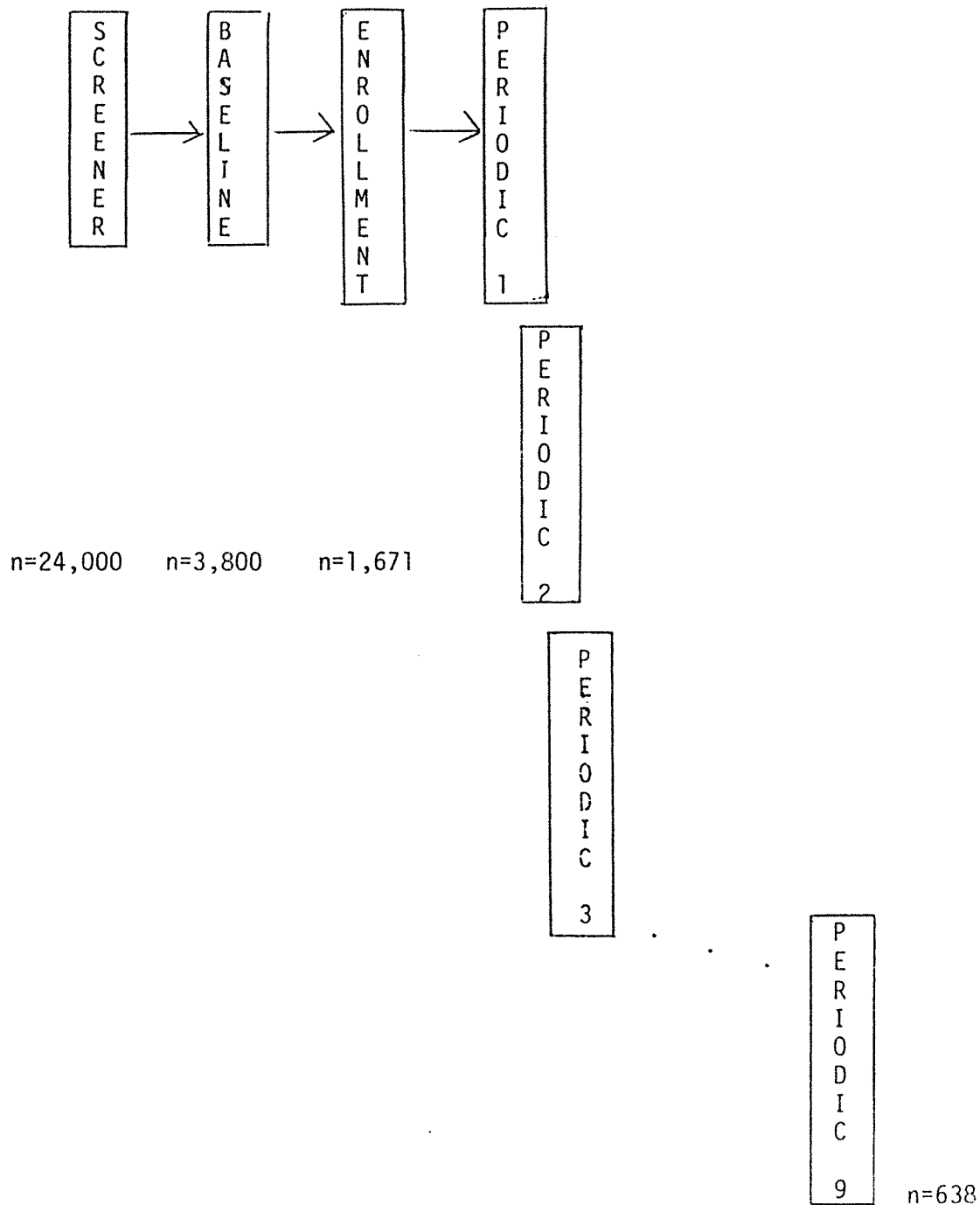
As an income maintenance experiment, MINCOME has several important features which serve to distinguish it. First, a panel was selected from a multiphase random sample of the Winnipeg population. The phases of selecting this panel can be seen below in Figure 1. The "screener" was selected on the basis of an average of 1969, 1970 and 1971 income and Manitoba Health Services Commission data on the demographic structure of households. Approximately 24,000 households were selected for this interview, which was the first stage in sample development and used a relatively short questionnaire, merely intended to identify potential eligibility. Of these screener households, 75% were in Winnipeg, 10% in rural areas, and 15% in Dauphin. The fact that random selection was used as the "screener" markedly differs MINCOME from similar U.S. experiments which used quota and area samplings. The scientific basis of the MINCOME panels implies that these data are inherently superior to those obtained in other experiments.

Second, a subset of the screener was selected for the "baseline" interview. Probabilities of entering the baseline were based on household type (double-headed, multiple earners; double-headed, single earners; single-headed; and single individuals) and income level. Households in each category were assigned a threshold income value (which increased with size of unit) at which the probability of selection became zero. No households with a head over 57 years in 1974 were selected. This truncation poses problems in hypothesis testing and estimation of model parameters discussed below.

Third, from the base, a subset of households were selected for enrolment in the experiment. Each household was assigned to the experimental cell with

Figure 1

Sample Development of Mincome

(Main Sample)


```
┌─────────┐      ┌─────────┐      ┌─────────┐      ┌─────────┐
│ S       │      │ B       │      │ E       │      │ P       │
│ C       │      │ A       │      │ N       │      │ E       │
│ R       │ ───► │ S       │ ───► │ R       │ ───► │ R       │
│ E       │      │ E       │      │ O       │      │ I       │
│ E       │      │ L       │      │ L       │      │ O       │
│ N       │      │ I       │      │ L       │      │ D       │
│ E       │      │ N       │      │ M       │      │ I       │
│ R       │      │ E       │      │ E       │      │ C       │
│         │      │         │      │ N       │      │         │
│         │      │         │      │ T       │      │ 1       │
└─────────┘      └─────────┘      └─────────┘      └─────────┘
```

```
                                            ┌─────────┐
                                            │ P       │
                                            │ E       │
                                            │ R       │
                                            │ I       │
                                            │ O       │
                                            │ D       │
                                            │ I       │
                                            │ C       │
                                            │         │
                                            │ 2       │
                                            └─────────┘
```

n=24,000   n=3,800   n=1,671

```
                                                    ┌─────────┐
                                                    │ P       │
                                                    │ E       │
                                                    │ R       │
                                                    │ I       │
                                                    │ O       │
                                                    │ D       │
                                                    │ I       │
                                                    │ C       │
                                                    │         │
                                                    │ 3       │
                                                    └─────────┘
```

```
                                                            ·  ·  ·
                                                                      ┌─────────┐
                                                                      │ P       │
                                                                      │ E       │
                                                                      │ R       │
                                                                      │ I       │
                                                                      │ O       │
                                                                      │ D       │
                                                                      │ I       │
                                                                      │ C       │
                                                                      │         │
                                                                      │ 9       │ n=638
                                                                      └─────────┘
```

probabilities also determined by income and size of unit. The design matrix

for the MINCOME experiment appears in Figure 2.

Figure 2

Experimental Design Matrix for MINCOME

Tax Offset

		.35	.50	.75
Income Support Guarantee	3800	1	3	6
	4800	2	4	7
	5800	*	5	8

*deleted due to expense.
(cell 9 is designated as the control group)

The offset tax rate was the rate at which income above the basic guarantee would be taxed. The income support levels were also adjusted upward by $600 in July 1975 to compensate for inflation and varied for different households. Figure 2 represents the support for a two-headed household of four people (2 dependents) at time of enrollment. One of the cells was purposely not tested, because of expense and cell 6 was allocated to cell 7 after one year because of sample size problems. The combination of low support and high tax was found to be too onerous and discouraged participation.

After the enrollment interview, nine periodics were conducted. These were comprehensive surveys administered to each participating family unit and probed job holding, job search, wage rates, sources of all income, disposition of assets, changes in family status, a variety of social indicators, etc. After the second periodic, it appeared as though attrition from the experiment was somewhat higher than similar experiments in the U.S. There appeared to be a danger that some of the experimental cells would have too small a sample to permit viable hypothesis testing and that attrition was systematically related to income and other independent variables. Also, a "hole" was discovered in the original sample frame where 16% of the eligible population may have been omitted from the screener survey. As a result, a supplementary sample was redrawn and a

sizeable group of families added to the experiment, which lags the main sample·
by about a year. Essentially, there are two experiments, basically of the same
design, one lagging behind the other. Although this complicates hypothesis test-
ing considerably should one wish to include both groups together it also opens
some interesting lines of study. The main sample started in the experiment
just as inflation began to accelerate in Canada; the second group entered when
inflation appeared, becoming a permanent feature.

3.0 METHODOLOGICAL ISSUES IN SOCIAL EXPERIMENTS

a. Previous Analysis

A number of authors have already identified some of the important
scientific issues in social experimentation. Lyall (1975) in particular
examines the New Jersey Income Maintenance experiment and cites the fol-
lowing major design problems.

i. Omission of Major Stratification Variables

Analysis of the New Jersey experiment has revealed, that the
simple two way design matrix with tax and guarantee on the two axes
is probably inappropriate. More important than the tax rate for
labour supply response is ethnicity. The income maintenance exper-
iments all failed to include this variable in the stratification of
the sample. The sample stratification undertaken at the screener
did not include ethnic group as a variable, and consequently one of
the most important variables for determining work incentive was not
included in the sample design. Of course, this and other variables
were monitored in the surveys and they were included as conditioning
or intervening variables in the actual estimation of labour supply
responses to a negative income tax. Nonetheless, much of the effort
in producing an optimal sample was fruitless since stratification
should have included ethnicity to improve the efficiency of subse-
quent analysis.

ii. <u>Narrow or Inappropriate Policy Spaces</u>

Analysis of most of the income maintenance experiments has shown that policy spaces are frequently too narrow. One must include dramatic differences in the guarantee levels to produce an appreciable alteration in work effort. Although this is a valid point from the point of view of the scientist, social experiments are constrained politically and ethically from incorporating extreme levels of treatments in the experimental designs. There is little point in producing variations which either have little chance of finding their way in policy variation or which may confound the experiment ethically. A negative income tax experiment, which, in the interests of scientific investigation allowed some of the participants to languish at the poverty level while others were made better off than the average family in society, is unlikely to find much political support. Certainly it may be questioned on ethical grounds.

iii. <u>Unanticipated Interactions within the Experimental Matrix</u>

All negative income tax experiments have shown that the vast majority of participants had little idea of the nature of the experiment. They operated from the after tax "take-home" pay perspective and were insentive to the variations in their "netback" from different cells in the matrix. It is interesting to note that in the MINCOME experiment, some participants were very sophisticated, and were able to compute exactly when it paid them to move out of the experiment and to adopt traditional welfare. In general, however the vast majority of participants were not responsive to the subtleties so carefully incorporated into the experiment.

iv. <u>Truncation and Attrition</u>

Lyall also gives brief consideration to the problems posed by attrition, refusal and truncation for the New Jersey experiment and

notes that "both the administrative costs of data collection and storage were underestimated" and that "truncating the sample on a dimension linked to the plan parameters produces a nonrandom allocation of various other family characteristics .... which may drastically constrain analysis of the final data." Her analysis of these problems is sparse however, primarily because in 1975, econometricians were still attempting to resolve the problems produced by very sophsiticated experimental designs. These specific issues are now examined in some greater detail.

b.   Truncation

All the negative income tax experiments had big but finite budgets. Accordingly, a great deal of attention was paid to a-priori design that would permit the evaluative exercise to proceed within reasonable cost. These experiments were primarily interested in a sub-population, namely the "working and not-working poor." The experimental designers were also sensitive to the problems of including groups who either could not complete the relatively onerous monitoring exercises, or whose decision making might confound the outcomes. For example, those with little ability in English and who were about to retire were not included in the experiments. All the data sets emerging from these social experiments are truncated.

As long as truncation can assume to be nonrandomly related to program parameters it can be incorporated into statistical analysis without much problem, however in none of the negative income tax experiments was this the case. For example, by eliminating those with little English, one would also diminish the size of the groups below a certain income level. More importantly, truncation also applied to income levels. Clearly, families which earn in excess of $29,000 (the current Manitoba average) would likely not be included in a negative income tax program,

but what about the group between $12,000 (assuming this is the cutoff chosen) and $29,000? Coupled with an income truncation rule was a cutoff pertaining to number of hours worked in a year. These two selection rules (income and hours worked) together produce a bias against harder working families near the income cutoff being included in the experiment and this may eventually bias labour supply response estimates downward.

In practice, this problem of truncation was handled by selecting relatively high cutoff levels, and by recently developed statistical procedures

c. <u>Attrition</u>

More important is the problem of attrition, especially since the degree of abandonment was initially higher than the MINCOME designers expected. Indeed, the early attrition in MINCOME was so high, especially in the least advantageous cell (#6) that this led, in part, to the decision to construct the supplementary sample.
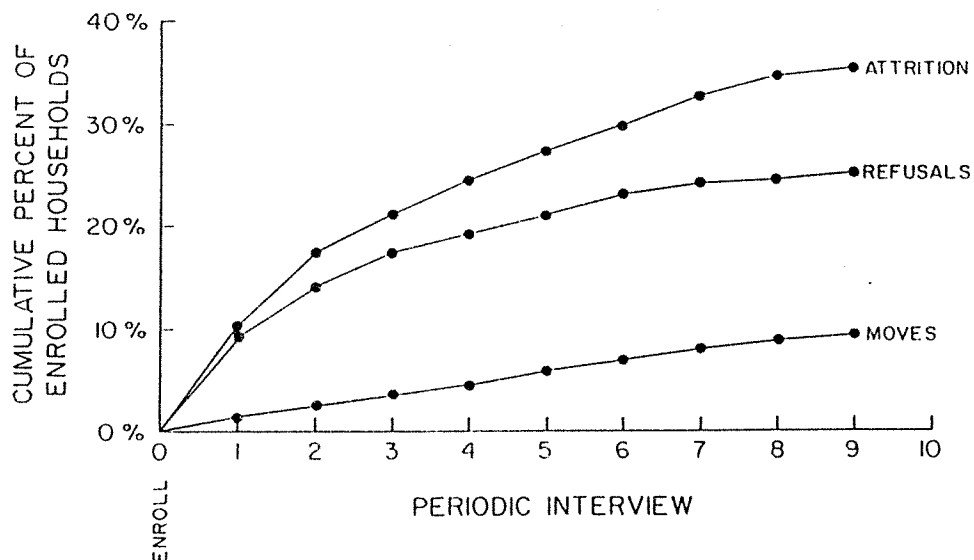
Attrition can be decomposed into two broad catagories - refusals and moves. In the former, all experiments undertook comprehensive analyses of the reasons why families either refused to be interviewed at the Baseline survey, who refused to enroll or who at some point in the experiment

decided to withdraw. Table 1 below presents the general catagories of
reasons why families left the experiment after initially enrolling .

Table 1

Reasons for Refusal after Enrolment (Winnipeg, Main Sample)

Reason Cited	Frequency	Percent
Don't like interviews--too personal	12	8.3
Just not interested in Mincome	11	7.6
No time to do reports	10	6.9
Not enough money received for participation	9	6.2
Forms ask too much information	5	3.4
New spouse refuses to cooperate	3	2.1
Personal reasons	3	2.1
Sibling refusal (refusal by a non-head)	1	.7
Family problems	1	.7
IRF too complicated	1	.7
Don't trust Mincome	1	.7
No reason given	88	60.6
TOTAL	145	100.0%

Source: Kurz, M. An Evaluation of the Experimental Sample of MINCOME
Manitoba Technical Report No. 5, Manitoba Basic Annual Income
Experiment 1979.

Most significant are the 60 percent who did not give a specific reason
for dropping from the experiment after they had received at least one pay-
ment. It may be conjectured that most of these participants left because
they found the experience too intrusive and burdensome, but were too polite
to indicate that to the interviewer.

The second aspect of attrition is the problem associated with moves
and in this aspect Manitoba was perhaps a problematic site, although this
note could be accurately forseen in the design of the experiment. At this

time (1974) stagnation was endemic to the province, and compared to western provinces little prospect existed that the future would become very much more hopeful. This induced many, especially young families and individuals, to leave in search of work. MINCOME suffered disproportionately from this disturbance to the basic sample.

The levels of attrition experienced in the first year were far too high to endure and still expect that the experiment could survive. Analysis indicated that the intial sample undertaken to develop the frame for the screener could have omitted as much as 16% of the eligible population. Thus the need to develop a supplementary sample seemed overwhelming. In hindsight, there are two major questions emerging from this decision. First, there is no sound theoretical method to integrate the supplementary and main samples. The supplementary sample is not a replication of what was done at the Baseline, but rather more closely represents a sophisticated quota sample designed to "fill" in "holes" in coverage. As such, it is doubtful that the two samples can ever be properly integrated, although recent developments in "metaanalysis" do offer some prospect for such a unified analysis. Second, the attrition problem did stabilize very rapidly, and experience in the second and third years proved much better than in the first. Overall, attrition for MINCOME resembles the experience of all the other experiments in the U.S. The evaluation of the sample developed by Kurz (1979) concludes it is possible to infer experimental effect from the data, provided that advanced statistical procedures are employed. This judgment however, reflects the view of one who wishes to analyse the data in light of the original experimental objectives. In fact, the MINCOME will support a comprehensive range of other analyses.

d.   Sophistication Versus Policy Formulation

Throughout the above discussion, the basic issue is the extent to which an experiment must adhere to theoretically accepted preconceptions. MINCOME, as the last negative income tax experiment, attempted to develop

the most comprehensive design, and understandably the designers attempted
to build upon and extend the work developed in the U.S. The desire to
evaluate the social and economic implications of a negative income tax for
Canada, placed great pressure upon the designers to develop and implement
the experiment prior to obtaining full information on the implications of
many of the design questions which were then emerging from U.S. experience.
The most important of these was the degree to which MINCOME adhered to and
elaborated upon the sample allocation model developed by Watts and Conlisk.
It is beyond the scope of this paper to detail this procedure, but a
schematic representation is useful.

Consider that MINCOME had set a total interview budget constraint
of $4 million for the life of the experiment (without the supplementary
sample). The usual way to allocate the sample over the design matrix
(figure 2) is to randomly assign participants with equal probabilities to
each cell. However, there is some theoretical justification for the
view that certain participants (households) produce more dramatic res-
ponses if they are located in some cells rather than others. If one
is interested in obtaining the greatest policy response for a given
budgetary expenditure, it makes some sense to allocate these responsive
household types to the cells where it was expected they would produce the
most useful information. The technique consists of hypthesizing a theor-
etical response surface (hours worked for example) which is a function of
the experimental variables (guarantee and tax) plus a number of independent
variables, and maximizing this function with respect to the total cost
constraint (interviewing and administration of the payments), using an in-
teger programming procedure. Provided all the relevant independent vari-
ables are included in the response surface, this procedure will produce
an allocation into the experimental matrix which provides a much greater
degree of response (policy) variation for a given cost.

Unfortunately, such an allocation will also produce an experimental design in which there are unequal probabilities for different participants within the experimental matrix. The usual analysis of variance procedures are not appropriate, since it is no longer possible to easily separate the effects of varying the policy variables and the affects of allocating the sample. Only recently have procedures been proposed to deal with this form of disturbance.

One might well ask why these experiments proceeded to employ a sample and experimental design for which there was no known analysis procedure. Considerable intellectual blood was shed over this issue, and each of the negative income tax experiments spent significant resources in attempting to resolve the problem. In general, those who advocated the new sample allocation procedures won the day, but in retrospect, this was probably a costly victory. Certainly the a-priori theory indicated significant cost savings and enhanced experimental value from a non-standard sample design, however the delayed analysis and the eventual need for more resources to complete the analysis has undoubtedly contributed to a mistrust of large scale experiments, especially those under the control of academics.

In retrospect, this issue of sophisticated design is more important than merely whether or not one particular group was right in terms of the theory; it relates fundamentally to the question of in whose interest these experiments were conducted. Economists largely determined the course of the experiments, and the experimental designers had in mind much more than merely whether the negative income tax was a sensible policy, they were much more concerned with academic kudos. It is not surprising that these experiments became overly complex. Occam's razor was set aside in the interests of undertaking a state-of-the-art design, capable of ensuring academic reward.

Although policy designers within governments expressed some concern

in this approach, it was only when MINCOME went over budget and more money
was requested that control was imposed upon the exercise. By then, it was
late in the day, and little could be done to modify the complexity.

This may sound too hard on academics, it probably is. Throughout
these experiments, and certainly this appears to have been true for MINCOME,
was a tension between the academics whose objectives were basically theor-
etical and policy makers whose needs were quite practical. No mechanism
to bridge and unify these desparate interests emerged and the gulf widened
to the point when fiscal problems forced the governments to intervene and
fundamentally change the nature of the program from a research endeavor
to an archival exercise. This problem is not unique to MINCOME but appears
to permeate policy design which involves teams composed of academics and
policy makers whose objectives diverge in important and fundamental aspects.

From the vantage, and comparative safety, of 1983 it is easy to state
that the overly complicated designs were poorly conceived. This may appear
to argue for simple approaches; there is no guarantee that these would have
produced the conclusions required in policy design. To have approached the
problem with a simpler framework and failed to produce the information
needed would have ellicited arguments in favour of more complicated ex-
periments. One cannot simply conclude on the basis of the failure of the
MINCOME experiment to conclude on time and within budget that its theor-
etical framework was fundamentally flawed. However, when administrative
issues are also considered, there emerges a persuasive case that MINCOME,
and other negative income tax experiments suffered from an over-sophisti-
cation which compromise the basic goals of the policy evaluation.

4.0 ADMINISTRATIVE AND LOGISTICAL ISSUES IN MINCOME

From the schematic of the experiment (see figure 1) it is apparent that a
very complex and arduous program monitoring exercise was underway. At its peak
of operation, MINCOME employed in excess of 230 people composed of four general

teams. An administrative staff administered and audited the payments made to each family, and accounted for almost half of the entire $17 million budget. A second and largest team was the field operations (interviewing) staff whose functions were to collect and check the data prior to data entry. A third group was concerned primarily with the experimental design and research analysis on the data. Finally, responding to the needs of the other three were a computer programming staff.

These four groups were coordinated through an executive director and a board composed of the directors of each of the four main teams. In reality, the research director assumed the role of deputy director, as the other three team leaders were primarily operational staff concerned with day to day functions.

The desparate nature of these four functions, which are generically representative of all major policy experiments, implied that enormous management pressure was placed upon the executive director and management board. To expect an individual and a small management board could coordinate the diverse and complex nature of such an exercise is quite unreasonable. Compounding this inherent structural complexity was a bureaucratic frontier. MINCOME was a separate legal entity empowered to disburse considerable sums of money and to prepare tax returns on behalf of the participants. It responded politically to both the Manitoba and federal governments, although eventually the influence of the province dwindled.

Of the four major groups within the operation, the administration of payments functioned without major problems, aside from the few expected cases of fraud and happenstance one would expect from the need to disburse regular payments to some 1000 participants over the entire province. Indeed, the best indication of the comparative ease with which this group functioned is that the payments data base was by far the easiest to access by the post experimental team at the Institute for Social and Economic Research. The need to produce a comprehensive and timely audit trail for continued funding from the two levels

of government ensured this task was well executed.

Somewhat more problematic were the field operations. This has little to
do with the actual mechanics of interviewing, since the usual problems associated
with survey research field operations occurred, if somewhat magnified by the
scale of the exercise. Rather, the sophisticated structure of the experiment
produced extreme complication in the administration of the surveys. In turn
this complication led MINCOME to adopt some illadvised "high tech" innovations
which failed, and this upset the critical path.

From figure 1, it may appear that the monitoring was to be accomplished
through a series of quasi-cross sectional surveys, separated in time by at least
a day. In fact, the surveys overlapped such that at any given point after the
Baseline was completed there were always two periodics in the field and at
times three surveys were underway. This can be best appreciated by considering
the frequency distribution of interviews, which is shown below in figure 4. The
horizontal axes represents weeks after the beginning of the Baseline (January
1, 1974) and shows surveys 4, 5 and 6 (Periodics 2, 3 and 4). The field op-
erations staff had to track when participants were interviewed, and scheduled
them into their next interview.

MINCOME and the other negative income tax experiments employed a wave sam-
pling procedure which allocated interviews within the periodics. Respondents
were positioned within the periodic to produce an average of four months be-
tween interviews. In other words, households who were surveyed early in a
periodic would always be positioned early in all subsequent surveys. The
logistics of surveying some 1000 households in a panel study necessitated
this approach, however, this technique results in data of a peculiar variety.
It is not correct to view the experiment as thirteen contiguous cross-sections.
Analysis of a pre-post test variety where the Baseline is compared to the last
survey is not feasible. For example, some households, interviewed early in
the periodic would report earning over a six week period; others would report

Baseline

Table 4

Enrolment

Periodic 1

Respondents per Week

Elapsed Time Since Baseline
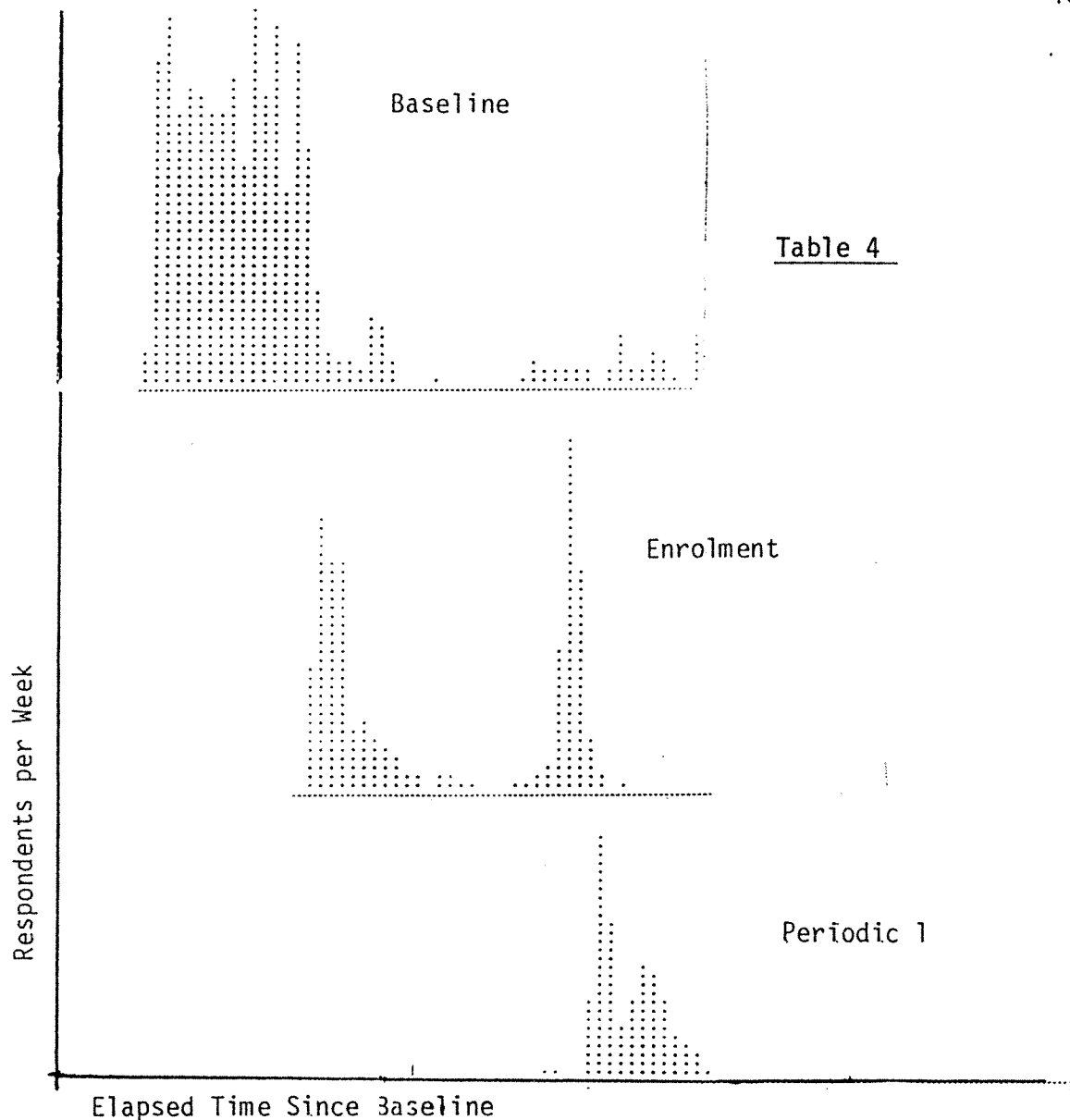
the same variable over a fifteen week period. Many variables must be averaged to an annualized equivalent. Also, in many cases, data from subsequent surveys must be used to construct a particular cross-sectional value, since not all questions were asked on each interview.

The actual field operations diverged from plan soon after the experiment began. The interviews took longer than anticipated and the periodics began to "stack" one another. Complicating this was a problem with data entry and verification which set back analysis to the point where cost overruns eliminated much chance of undertaking research within the time frame originally allotted to the experiment. Initially, the MINCOME team planned to use the university

computer on a time sharing basis, however political interest intervened and dictated that, in the interest of regional industrial development, MINCOME would be required to use the fledgling government computer service bureau. Aside from its relative inexperience in dealing with large scale data processing requirements, this service bureau was really set up to process accounts receivable and other financial data. It had modest capabilities to support the needs of a large social science projects. In the face of this, MINCOME elected to proceed using a mini-computer. All of the data entry, verification and eventual data base design was accomplished on a PDP-11. At no time was the computer programming group ever able to inspect an entire periodic let alone access to the entire accumulated data.

Another, somewhat more serious problem was, an abortive experiment with optical scanning. This involved character reading (as opposed to the much more developed field scanning techniques) and interviewers were instructed in the appropriate handwriting techniques required to allow the optical scanner to extract information from the questionnaires. Presumably, this would eliminate the lengthly and error prone process of data entry and verification. Unfortunately, after protracted efforts to operationalize the system, it was abandoned and the more traditional approach of manual questionnaire checking, data entry and validation resumed. At this point the Baseline had been completed and the enrolment interviews were well underway. This delay implied that the data entry operations always lagged the interviewing process by a relatively long period. Questionnaires were not checked immediately upon receipt and cross-sectional validity checks were not undertaken at the point of data entry until comparatively late in the experiment.

The fact that a small computer was used to undertake all aspects of MINCOME operations (including payments administration, interview scheduling as well as data entry) produced extreme strain on the computer staff. At no point in the experiment can it be said that computer operations were satisfactory and cost-

A final general problem with respect to logistics was the problem in meld-
ing a variety of objectives and disciplines into a coherent operating unit.
Each of the administrative units (payments, research, field operations and com-
puter systems) all functioned within an internally consistent set of techniques
and procedures. Unfortunately, the bridge between these separate functions was
only imperfectly constructed. Nowhere is this more apparent in the design of the
main data base to conserve the data for post-experimental analysis.

In 1976, the experiment underwent a fundamental change in direction.
Partly because of shift in the political views and partly because of cost over-
runs due to the need to conduct a supplementary sample and the sudden surge in
the Canadian inflation experience, MINCOME was converted from a policy oriented
experiment to an archival exercise. After 1976, MINCOME was expressly forbidden
to undertake any research and to concentrate solely upon data collection, pay-
ments administration and the construction of a post-experimental data base which
would permit research by a yet to be named research team. Many in MINCOME ex-
pected (hoped) that in some manner the essential research team would be retained
for this post-experimental analysis and the design of the data base reflected,
in part, this anticipation. Also, and perhaps more importantly, the selection
of the data base was conditioned by systems analysts who had primary experience
with large managerial data bases, as opposed to the information needs of policy
planners and social scientists. A state-of-the-art system was selected as the
host system into which the MINCOME information was to be configured. This
data base management system (DBMS) has had success with large administrative
files (such as hospital and corporate records) but its capabilities are limited
in presenting the "flat" analysis files required for policy evaluation employing
social science techniques. Very considerable programming is required to pro-
duce the data formats needed by the standard statistical software. Coupled with
the constrained computer at MINCOME and the haste which characterized the clos-

ing of the experiment in 1979, the eventual DBMS has proved quite unworkable. This administrative and logistical experience of MINCOME reflects the imposition of theoretical complexity upon a large sample which required regular and intricate monitoring. All the complexities of a cross-sectional survey exist, but are magnified exponentially by the requirements of a panel study. In some important ways, the theoretical requirements of MINCOME (and the other negative income tax experiments) exceeded the capabilities of social science and computing technologies exsisting in the mid seventies.

5.0 THE LESSONS FROM MINCOME

A number of very constructive lessons have been learned from the MINCOME and other negative tax experiments.

First, and most importantly is that despite the many errors and problems, the data collected in these surveys is remarkably robust. Analysis on the data using the Baseline indicates that it is indeed useful and capable of supporting a very wide range of social policy issues. Social scientists tend to be very purist in their approaches and often impose very unrealistic standards on data quality. Of course, had any of the errors been more significant, for example had attrition been over 50%, this statement might be wrong. The fact that the information is sound reflects in large measure the committment and efforts of individual personnel within the experiments who frequently extended themselves to preserve the basic integrity of the data collection and entry. The human factor, so often cited and the cause of problems, can also resolve all manner of logistical difficulty.

A second lesson that is also very clear is that methodological purity can often exceed technological capacity. In this, social scientists attempt to mimic the control found in advanced physical science laboratories, and fail to appreciate the multifaceted milieu in which all social experiments must operate.

During the life of MINCOME for example, the inflation rate rose dramatically and caused the imposition of wage and price controls, certainly a potentially very significant influence on work incentives and behaviour. Complex experiments with detailed policy parameters are unlikely to produce either theoretically successful results, or more importantly persuasive input to the policy formation process. The basic aim of policy design must supercede the needs of academics to produce research acceptable to a small group of their peers.

Related to this is the general problem of academic arrogance. So much of the negative income tax experimental experience reflects the self-assurance of theorists with little direct experience in the type of behaviour that would be expected in actual administration of an experiment. Of course, there are particular details about a given policy which may not be anticipated, such as the high rates of moves within the MINCOME experiment, however, ethnicity should not have been ommitted in the design of the original samples, especially for the New Jersey situation. A well designed pilot study, plus examination of the literature should have eliminated this error.

A third lesson is that the theoretical design cannot be divorced from the administrative and logistical features of the experiment. For example, if the basic objective is scientific policy design, then this objective should supercede the need to use the expenditures for industrial development. At some point the research objectives are seriously compromised, and to persist with the excercise ultimately causes social science and evaluation to be cast in a very poor light. Of course, the degree of committment required of a social experiment will induce many to linger in rear guard actions well past the point when sunk costs will never be recovered. As long as the professional social science community is willing to accept the imposition of non-evaluative policy goals, then the typical politician will assume that there is no damage to the basic integrity of the research.

It is intriguing to speculate whether academics should ever figure prominently in the overall direction of such large scale projects. On the one hand

their expertise is essential to the basic scientific integrity of the exercise; however, their longer term objectives are to obtain <u>academic</u> credit for the research. A civil servant might be expected to have a much finer grasp of the policy needs of the politicians, but may not be sensitive to the compromises which can upset the validity of the research. What is required is a very special type of researcher. On the one hand, they must gain self-esteem from managing a complex logistical task. Very few academics place much stock in the tasks of collecting high quality data; of all the social sciences, perhaps economics is the worst offender. The overall coordinator of a major social experiment must combine advanced training in the social sciences with political acumen. In large scale research projects which require computer processing and the design of elaborate retrieval systems, a comprehensive knowledge of systems analysis is needed, if for no other reason than to keep the programmers honest.

The above remarks are general and certainly do not constitute a criticism of the MINCOME personnel. Rather they are directed at an academic community which undervalues high quality data collection, and a civil service which is insensitive to the requirements of sound policy design.

Will large scale social experiments be attempted in the near future? I think not-despite the fact that the technology now exists which can alleviate many of the problems encountered in MINCOME. The political mood is far away from the heady bouyancy which characterized the optimism of the "great society". In time, as more modest experiments prove their worth in policy design we may witness a new era of policy design and the recent experience with MINCOME and other social experiments will prove invaluable.

REFERENCES

1. Campbell, P.T. (1971) "Methods for the Experimenting Society" American Psychological Association, Annual Meetings, 1971.

2. Institute for Social and Economic Research MINCOME User Manual (Version 20), Feb. 1983, University of Manitoba.

3. Kurz. M., "An Evaluation of the Experimental Sample of MINCOME Manitoba". ISER #(MMTP-6) Institute for Social and Economic Research, University of Manitoba.

4. Lyall, K. C. "Some Observations on Design Issues in Large-Scale Social Experiments", Sociological Methods and Research, Vol 4, #1, August 1975 pp. 54-76.

5. Rasmussen, E., et al., "MINCOME Field Operations and Interviewing Techniques", ISER #(MMTP-12), Institute for Social and Economic Research, University of Manitoba.